(1)

Research Product 84-19

HUMAN RESOURCES TEST AND EVALUATION SYSTEM (HRTES) VOLUME 1: TEST PROCEDURES

ARI Field Unit at Fort Hood, Texas
Systems Research Laboratory

August 1984

U.S. ARMY RESEARCH INSTITUTE for the BEHAVIORAL and SOCIAL SCIENCES

U. S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

L. NEALE COSBY
Colonel, IN
Commander

Research accomplished under contract
for the Department of the Army

Perceptronics, Inc.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ARI Research Product 84-19	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) HUMAN RESOURCES TEST AND EVALUATION SYSTEM (HRTES) VOLUME 1: TEST PROCEDURES		5. TYPE OF REPORT & PERIOD COVERED Final - May 1982 6. PERFORMING ORG. REPORT NUMBER None
7. AUTHOR(s) Jonathan D. Kaplan (ARI), Norman D. Schwalm & William H. Crooks (Perceptronics)		8. CONTRACT OR GRANT NUMBER(s) DAHC-19-77-C-0055
9. PERFORMING ORGANIZATION NAME AND ADDRESS Perceptronics 6271 Variel Avenue Woodland Hills, CA 91367		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Program Element 2Q263739A793 Task 3229 Work Unit 53
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue, Alexandria, VA 22333-5600		12. REPORT DATE August 1984 13. NUMBER OF PAGES 69
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) --		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE --

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

--

18. SUPPLEMENTARY NOTES

Volume 2 of this report serves as an appendix and appears as RP 84-20.

Keywords:

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Human performance
Human factors
Human factors engineering
Operational testing
Test and evaluation
Human resources
Training and evaluation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Human Resources Test and Evaluation System (HRTES) is a set of procedures designed to assist a test planner in evaluating the operator and maintainer performance in an operational test of an Army system. Specifically, HRTES provides guidance for (1) identifying the critical aspects of human performance to be included in the operational test, (2) evaluating that performance, and (3) analyzing the cause(s) of any inadequate human performance. Guidance is also provided for identifying the contributions of (Continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A1

ARI Research Product 84-19

20. (Continued)

training, personnel selection, and human factors engineering to overall human performance. This volume consists of a series of brief steps stating the tasks to be performed in a human resources evaluation of an Army system. The procedures involved in each step are described and the rationale is given when appropriate.

Research Product 84-19

HUMAN RESOURCES TEST AND EVALUATION SYSTEM (HRTES) VOLUME 1: TEST PROCEDURES

Jonathan D. Kaplan
Army Research Institute

Norman D. Schwalm and William H. Crooks
Perceptronics, Inc.

Submitted by
George M. Gividen, Chief
ARI Field Unit at Fort Hood, Texas

Approved as technically adequate
and submitted for publication by
Jerrold M. Levine, Director
Systems Research Laboratory

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333

Office, Deputy Chief of Staff for Personnel
Department of the Army

August 1984

Army Project Number
2Q263739A793

Human Factors in Training & Operational Effectiveness

ARI Research Reports and Technical Reports are intended for sponsors of R&D tasks and for other research and military agencies. Any findings ready for implementation at the time of publication are presented in the last part of the Brief. Upon completion of a major phase of the task, formal recommendations for official action normally are conveyed to appropriate military agencies by briefing or Disposition Form.

FOREWORD

Because of their complexity, weapons and other systems being acquired by the Army will place heavy demands on operator and maintenance personnel. To avoid costly failures, the human resources aspects of these future systems must be fully evaluated early in their development cycles, so that problems can be corrected prior to full-scale production. However, the Army does not have enough personnel with human factors expertise to staff all of the system tests conducted each year. Personnel with little or no training or experience in the human factors area must often be assigned to conduct these evaluations. The Human Resources Test and Evaluation System (HRTES) was designed to meet the need for a guidance document to aid the "typical" test officer in planning and conducting human resources evaluations of proposed Army equipment.

EDGAR M. JOHNSON
Technical Director

HUMAN RESOURCES TEST AND EVALUATION SYSTEM (HRTES)
VOLUME 1: TEST PROCEDURES

EXECUTIVE SUMMARY

Requirement:

Weapons and other systems being acquired by the Army are becoming increasingly complex and costly, and place ever increasing demands on operator and maintenance personnel. Recent data suggest that these personnel are responsible for over one-half of the failures of major systems. Therefore, it is imperative that the human resources aspects of future systems be fully evaluated early in the development cycle, and that problems be corrected prior to full-scale production. However, the Army does not have adequate numbers of personnel with human factors expertise to man all of the system tests conducted each year. As a result, personnel with little or no training or experience in the human factors area must often be assigned to conduct human factors evaluations. Therefore, there is an obvious need for guidance documents to aid the "typical" test officer plan and conduct human resources evaluations of Army equipment. The Human Resources Test and Evaluation System (HRTES) was designed to meet this need.

Procedure:

In developing HRTES, it was assumed that the primary purpose of tests and evaluations was to determine whether the tested systems were able to satisfy the requirements for which they were developed. Given this assumption, procedures were developed to focus first on identifying those activities or functions a system must perform. Since the emphasis in HRTES was to be on the human components of a system, procedures were then developed to identify those human activities which must be performed for the system as a whole to perform its functions. Next, procedures were developed to determine what aspects of human performance had to be measured. Finally, procedures were developed for analyzing the cause(s) of any inadequate performance. This latter guidance was designed to aid the test officer in identifying the contributions of training, personnel selection, and human factors engineering to overall human performance.

Volume 1 of HRTES, titled TEST PROCEDURES, is the primary guidance document. It describes the steps to be taken in performing each of the major tasks. Volume 2, titled SUPPLEMENT, contains detailed descriptions of a number of the test procedures and methods. Thus, the supplement can be considered to be an appendix to the Test Procedures volume.

HRTES

Human Resources Test and Evaluation System

CONTENTS

1. INTRODUCTION

1.1 Overview

This Human Resources Test and Evaluation System (HRTES) is a set of procedures designed to assist a test planner in evaluating the operator and maintainer performance in an operational test of an Army system. Specifically, HRTES provides guidance for (1) identifying the critical aspects of human performance to be included in the operational test, (2) evaluating that performance, and (3) analyzing the cause(s) of any inadequate human performance. Guidance is also provided for identifying the contributions of training, personnel selection, and human factors engineering to overall human performance.

This chapter includes a brief description of the contents of HRTES and a summary of the procedures that can be used to identify and evaluate human performance in an Operational Test. The chapter also describes the products that will be developed when HRTES is used, and the relationship of these HRTES products to the other documents that are produced during the course of planning and conducting operational tests. This chapter concludes with a description of the organization of these two HRTES volumes.

1.2 Contents of HRTES

The human factors of a system in HRTES include such "supporting activities" as training, personnel selection, human factors engineering, and safety and health.

For purposes of evaluating the human factors of a system, it is useful to focus first on evaluating human performance, and then to focus on determining the contributions of the "supporting activities" to any instances

of inadequate human performance that are identified during the course of a test. From this perspective, the human performance in an Army system is all of the operator and maintainer actions that contribute to overall system performance. Thus, planning the human resource aspects on an operational test initially consists of identifying the operator and maintainer performance that will be measured during the OT. Subsequent activities in the test planning focus on evaluating the human performance and analyzing the contribution of the supporting human factors activities to instances of inadequate human performance.

Because operator and maintainer performance is considered to be a subset of overall system performance, the first step in identifying operator and maintainer performance is to identify the overall system performance that is to be evaluated in the operational test. In HRTES, the overall system performance to be evaluated is identified on the basis of the test issues that are specified for the operational test of the Army system. Based on the test issues, the associated scope, and the criteria, the HRTES user then identifies the operator and maintainer tasks to be evaluated in the OT. In addition to providing guidance for identifying the tasks, HRTES also provides methods for identifying the measures to be used during the OT for assessing the human performance.

In addition to measures of human performance, the operational test will include many direct measures of training, human factors engineering, safety and instances of accidents and critical incidents. The data from these measurements will be important during evaluation to help identify and analyze inadequate human performance. HRTES provides guidance to the test planner for identifying these additional human factors measures that must be included in the test plan.

After operational issues are measured, they are evaluated by comparison with criteria. After the human performance is measured during the operational test, that performance required for inadequately performed issues is evaluated. For those instances of inadequate human performance, a process of analysis is used to identify the most likely cause(s). Such inadequate human performance can usually be attributed to deficiencies in the supporting activities of training, personnel selection, and human factors engineering. A key problem in analysis is to determine that the deficient performance is indeed inadequate human performance (rather than poor hardware or software performance) and then to determine which of the several supporting activities contributed to the deficient performance. HRTES provides guidance not only for evaluating all of the tested human performance, but also for diagnosing the probable sources of inadequate performance.

1.3 Using HRTES

HRTES is intended to be used by persons having responsibility for the human resource aspects of OT test planning and evaluation. HRTES is organized as a series of steps (or tasks) that are performed in planning and evaluating human factors. In performing these steps, the HRTES user will produce a number of products (usually portions of documents) that can be used directly as part of the documents prepared in the operational tests (i.e., IEP, OTP, TDP, DTP, TR, IER).

1.3.1 HRTES - Test Procedures. The Test Procedures volume of HRTES provides the primary guidance in planning and evaluating human performance for the operational test. This Test Procedures volume consists of a series of brief steps stating the tasks to be performed. Following the statement of each step is a brief description that explains the procedure and gives a rationale where appropriate. For those steps that may involve a number of substeps, or where several alternative methods are available for performing the step, the reader is referred to the HRTES - Supplement volume.

Each step in the Test Procedures volume is highlighted by a surrounding box. This identifies the specific tasks to be performed by the HRTES user. It is also intended to assist a user who has previously used HRTES, since the explanations will probably not be necessary on subsequent uses.

The steps in the Test Procedures volume are listed in sequential order according to a typical sequence of test planning. However, the steps can be performed in any sequence that is appropriate for the user.

1.3.2 HRTES - Supplement. The Supplement volume includes the detailed descriptions of a number of test procedures and methods. These descriptions provide further explanation of the steps contained in the Test Procedures volume. The procedures in the supplement are referenced from the Test Procedures; thus the Supplement can be considered to be an appendix to the Test Procedures. The detailed procedures in the Supplement are used whenever additional detail is required beyond that given in the Test Procedures volume, or whenever an alternative method for performing the step is needed.

2. IDENTIFYING TEST ISSUES

2.1 Overview

Typically, the first step in planning a test is to identify the major questions that are to be answered by the test. In the case of operational tests, these major questions are given as test issues. The Human Resource Test and Evaluation System (HRTES) procedures are designed specifically to aid the test planner in identifying the human factors of an Army system, i.e., those parts of an Army system that affect or are affected by human performance. Therefore, it is useful to review the test issues as a source for identifying those human performance questions that must be answered about the weapon system. Associated with an issue is a description or specific list of conditions (called the "scope") under which the issue should be tested. These conditions should define the range of variables that affect human performance and that should be considered in the test. The starting points for planning the operational testing of human factors components of Army systems are to determine that 1) the issues included will require the testing of important aspects of human performance, and 2) the scope included will specify the variables that may affect human performance. While the test issues and their scope constitute the questions to be asked about the Army system in the operational test, the criteria for the issues focus on the answers that must be obtained. The criteria define the general form of measurements to be taken and the acceptable range of variation in each measurement. Thus, for human performance concerns, the criteria of the test issues will have direct implications for the types of human performance measures that must be planned for the test.

This chapter provides guidelines for identifying the test issues, scope, and criteria that will have human resources implications for test planning. Guidelines are also provided for developing issues, scopes, and criteria

for those instances where they do not exist or where they are not adequate for developing the human performance aspects of the test plan.

2.2 Identifying Test Issues

The first step in planning the human performance portions of an operational test is to identify the test issues that will have human performance implications. In some cases, the documentation associated with the Army system will be sufficiently thorough to include all of the issues with important human performance implications. In these cases, the first task is primarily a matter of selecting the issues from these documents and verifying that they will include the measurement of important human performance aspects of the system. In other cases, some, but not all, of the human performance related issues will be included in the system documents. In these latter cases, the task is to develop the additional issues that will be necessary to test all of the important human performance implications.

2.2.1 Selecting Test Issues

1. IDENTIFY THE ISSUES THAT ARE APPLICABLE TO THE SYSTEM BEING TESTED.

Depending on the stage of Army system development, some or most of the test issues may already be given. The task at this point in the test planning is to determine that all issues having potential human performance implications have been identified. The relevant test issues are usually contained in the documents associated with the development of the system, including:

ROC/LR - Required Operational Capability, or Letter Requirement

MENS - Mission Element Needs Statement

IEP - Independent Evaluation Plan

OAP - Outline Acquisition Plan

LSA - Logistic Support Analysis

and others.

2. REVIEW THE FULL SET OF ISSUES TO IDENTIFY ALL ISSUES THAT HAVE HUMAN PERFORMANCE IMPLICATIONS.

A distinction can be made between two classes of test issues, namely: 1) "system operability" issues that directly examine the performance of the system, and 2) "system supportability" issues that examine the various activities that contribute indirectly to system performance. Although they may be tested differently, both classes of test issues have implications for human performance that can be tested. The following categories of test issues all have implications for human performance in the Army system. This list should be used as a general checklist to verify that all potential issues having human performance implications have been identified for the system.

CATEGORIES OF ISSUES
WITH HUMAN PERFORMANCE IMPLICATIONS

SYSTEM OPERABILITY

1. Mission Performance
2. Survivability/Vulnerability
3. Reliability/Availability/ Maintainability (RAM)
4. Doctrine
5. Transportability
6. Interoperability

SYSTEM SUPPORTABILITY

1. Training
2. Personnel Selection
3. Human Factors Engineering
4. Safety/Health

2.2.2 Developing Test Issues

Note: Skip this section if the issues with human performance implications have been identified adequately.

1. DEVELOP ANY REMAINING TEST ISSUES THAT HAVE HUMAN PERFORMANCE IMPLICATIONS.

In some cases, important categories of test issues with human performance implications may not have been included in the documents associated with the Army system development. To assure that the human performance aspects of the Army system are treated adequately in the operational test, these additional test issues must be developed.

Procedures for developing test issues are given in HRTES - Supplement, Section S 2.1.

2.3 Identifying the Scope of Test Issues

The scope of an issue consists of the conditions under which the issue will be tested. These conditions often suggest the nature of the human performance that can be examined in the operational test. Thus, identifying the scope of the test issues is a critical step in planning for the human performance aspects of the OT. As in the case of identifying the test issues, the documents associated with the Army system may include

all of the conditions that will be important for testing the human performance implications. The task at this point will be to select these conditions and to verify that all of the important conditions that may affect human performance have been included. In those cases where some important test conditions have not been included, these conditions should be developed for the scope of the issues.

2.3.1 Selecting Scope

1. DETERMINE THE CATEGORIES OF SCOPE THAT ARE APPLICABLE TO THE ARMY SYSTEM BEING TESTED AND TO THE PLAN BEING DEVELOPED.

The task at this point is to identify all of the conditions that potentially may apply to the human performance that will be tested for this system. The following list of categories may be used as a checklist to verify that all test conditions having potential human performance implications have been identified:

CATEGORIES OF SCOPE WITH HUMAN PERFORMANCE IMPLICATIONS

WEATHER:
- Illumination
- Temperature
- Precipitation
- Wind
- Humidity

TARGET:
- Type
- Number
- Location
- Speed
- Direction of Motion
- Concealment

TERRAIN:
- Ground Slope
- Ground Surface
- Ground and Water Surface
- Obstacles

PERSONNEL:
- Workload
- Duration of Preceding Work
- Protective Gear
- Physical Strength
- Perceptual Ability
- Experience
- Aptitudes
- Physical Size

TRAINING:
- Institution
- Latency
- Team vs. Individual

OPERATIONAL:
- Crew
- Hardware
- Information Inputs

TACTICS:
- Number of Systems Employed
- Speed
- Location
- Direction of Motion
- Concealment
- Crew Protection
- Amount of Automatic Functioning
- System Workload

Some of the preceding categories of scope may not be relevant to the procedures of your organization for selecting test conditions. If this is the case, and if several significant categories of human performance would thus be overlooked, it may be useful to coordinate the selection of test conditions with the appropriate organizations.

> 2. REVIEW THE SCOPE ASSOCIATED WITH EACH TEST ISSUE AND DETERMINE THAT ALL CONDITIONS WITH IMPORTANT HUMAN PERFORMANCE IMPLICATIONS HAVE BEEN INCLUDED.

Not all categories of scope that were identified as having human performance implications should necessarily be associated with all test issues. The task at this point is to determine that the previously identified categories of conditions are included with one or more test issues. Once you have decided which scope categories are applicable to human factors considerations of the system, examine the scope listed in the

existing system documents. Determine which applicable categories are missing. Specific conditions must then be developed for these scope categories.

2.3.2 Developing Scope

Note: Skip this section if the issues and scopes with human performance implications have been identified adequately.

1. DEVELOP TEST SCOPE(S) TO INCLUDE THE REMAINING CONDITIONS THAT HAVE IMPORTANT HUMAN PERFORMANCE IMPLICATIONS.

In some cases, conditions that have been identified as having important human performance implications may not have been included in the scope of the test issues. The task at this point will be 1) to identify those conditions that have not been included as part of the scope of one or more issues, 2) to determine which test issue(s) should include these remaining conditions, and 3) to expand the scope of these issues to include these remaining conditions.

Procedures for developing the scope of issues are given in HRTES - Supplement, Section S 2.2.

2.4 Identifying Criteria of Test Issues

An important step in test planning is to identify the criterion for each of the test issues, since the criterion directly implies the form of the measurements to be taken during the test and the range of variation that must be measured. Here too, the documents associated with the Army system may already list the criteria for the test issues, or some criteria

may not have been specified for some issues. In this latter case, the criteria must be identified as a basis for establishing the measurements that will be required during the test.

2.4.1 Identifying Issues with Multiple Criteria

1. FOR EACH ISSUE, DETERMINE IF ANY OF THE CONDITIONS IN THE SCOPE WOULD BE MUTUALLY EXCLUSIVE IN A SINGLE PERFORMANCE TRIAL.

The first step in identifying the criterion of a test issue is to determine if any conditions of that issue are mutually exclusive. Conditions are mutually exclusive if they cannot be represented at the same time in a single trial that tests an issue (e.g., "day" and "night" conditions cannot exist simultaneously in a single trial of a test for "adequacy of firing"). Mutually exclusive conditions imply that different criteria may be appropriate for the exclusive conditions.

2. FOR ISSUES WITH MUTUALLY EXCLUSIVE CONDITIONS, DETERMINE THE COMBINATIONS OF CONDITIONS THAT WILL BE PERFORMED TOGETHER IN A SINGLE PERFORMANCE TRIAL.

Note: Skip this step for those issues with non-exclusive conditions.

There are two general methods for determining the combinations of conditions that should be tested for an issue, as follows:

Method 1. Divide the scope of the issue into every possible combination of conditions.

EXAMPLES: 1) Day, Target range: 2000-3000 meters, target moving laterally, Ammunition: Sabot, Personnel wearing NBC gear, following 24 hours of continuous combat, etc.

2) Night, Target range: 2000-3000 meters, Target moving laterally, Ammunition: Sabot, Personnel wearing NBC gear, Following 24 hours of continuous combat, etc.

3) Day, 1000 meters, laterally moving target, Sabot, wearing NBC gear, after 24 hours of continuous combat.

4) Night, 1000 meters, etc.

Method 2. For a given issue, divide its scope into only the most significant combinations of conditions that you wish to test.

EXAMPLES: 1) Day, 2000 meters, etc.

2) Night, 1000 meters, etc.

(Note that combinations 2 and 3 in Method 1 have been left out since they have been considered less significant for testing).

If the scope of an issue contains a very small number of exclusive conditions, the first method for determining combinations of conditions is a reasonable one. If there are a substantial number of mutually exclusive conditions, the second method (selection of only significant combinations) may be more useful.

In determining which conditions to combine, the following guidelines may be helpful:

A) An issue should be tested under the combination of conditions that represents the situation normally expected in battlefield use.

B) An issue should also be tested under the combination of conditions that represents the worst case that can reasonably be expected in battlefield use.

3. RECORD THE ISSUE AND COMBINATIONS OF CONDITIONS THAT WILL APPLY TO THE ISSUE DURING THE TEST.

For those issues that have non-exclusive conditions, it is useful to represent the issue and its scope in a matrix as shown:

EXAMPLE: Matrix of an issue with non-exclusive conditions:

ISSUE / SCOPE	Day	Trgt Moving Laterally	Trgt Range--2000-3000M	Ammo - Sabot	Personnel - NBC Gear	Personnel - After 24hrs Continuous combat	. . .
1. Target Acquisition	x	x	x	x	x	x	
.							
.							
.							

For those issues with mutually exclusive conditions, the issue and scope may be represented in matrix form as follows:

EXAMPLE: Matrix of an issue with mutually exclusive conditions:

ISSUES \ SCOPE	Day	Night	Range--2000-3000M	Ammo-Sabot	. . .
1. Target Acquisition Set 1	x		x	x	
Set 2		x	x	x	
. . .					

2.4.2 Identifying Criterion of Test Issues

1. FOR EACH ISSUE, IDENTIFY THE CRITERION FOR ADEQUATE PERFORMANCE.

The criterion of an issue depends on the nature of the issue. In the case of "system operability" issues, the criterion should be specified in terms of system performance. Such performance should be specified in terms of acceptable accuracy and time of performance over an acceptable number of trials.

In the case of "system supportability" issues, the criterion should be stated in terms of the performance that is to be supported, or in terms relative to the supportability characteristics of a similar system.

A criterion for "system operability" issues should have the following characteristics to be useful in an OT"

A) A statement of the actual performance required by the issue.
EXAMPLE: Destroy 4 out of 5 ememy targets.

B) A definition of the minimum acceptable accuracy level for one trial.
EXAMPLE: Put at least 3 rounds within 1 meter of the center of mass of a given target.

C) A definition of the maximum acceptable amount of time permitted for one trial.
EXAMPLE: Put at least 3 rounds within 1 meter of 4 out of 5 targets within 20 seconds.

D) A definition of the criterion of performance for multiple trials within the constraints given for a single trial.
EXAMPLE: Put at least 3 rounds within 1 meter of 4 out of 5 targets within 30 seconds in 80 percent of the trials (or with an 80 percent probability of occurrence).

A criterion for a "system supportability" issue should contain a statement based on operational acceptability. EXAMPLE: Safety (or HFE, training, personnel selection, etc.) must be adequate to permit the system to operate within all its performance criteria without injuring personnel (or exceeding given parameters of training, personnel selection, etc.).

In those cases where the criteria for an issue have not been stated, or where the criteria do not meet the standards described above, the criteria must be developed.

2. FOR ISSUES WITH MUTUALLY EXCLUSIVE CONDITIONS, DETERMINE THE CRITERIA THAT ARE ASSOCIATED WITH EACH COMBINATION OF CONDITIONS.

Note: Skip this step for issues with non-exclusive conditions.

When an issue has more than one combination of conditions, it may not be clear (1) that the criterion applies to all combinations, or (2) which criteria apply to which combinations of conditions. Several steps should be performed at this point, namely:

A) Determine if the criterion applies to each combination of conditions. This may be done in two ways:

1) Determine if the criteria apply only to mutually exclusive conditions.

EXAMPLE: Issue to be performed in 10 seconds in the day, versus 20 seconds at night. In matrix form, this may be represented as:

ISSUES / SCOPE	Day	Night	Range--2000-3000M	Ammo-Sabot	CRITERION
1. Target Acquisition Set 1	x		x	x	10 sec.
Set 2		x	x	x	20 sec.
. . .					

2) Determine if the criterion applies to all combinations of conditions.

EXAMPLE: Issue to be performed in 10 seconds under all conditions. In matrix form, this may be represented as:

ISSUES / SCOPE	Day	Night	Range--2000-3000M	Ammo-Sabot	CRITERION
1. Target Acquisition					
Set 1	x		x	x	10 sec.
Set 2		x	x	x	
. . .					

B) Match the combination of conditions to the appropriate criterion. Occasionally, issues will appear with 1) mutually exclusive conditions, 2) only one criterion, and 3) with conditions to which that single criterion may or may not apply. There are three questions that should be answered:

1) Did the organization that prepared the criterion intend that it apply to all combinations of conditions?

If the answer is "yes," questions (2) and (3) are not relevant. If the answer is "no," questions (2) and (3) must be answered.

2) To which of the combinations of conditions did the organization that prepared the criterion intend the criterion to apply?

3) What is the issue criterion for the remaining combination of conditions?

2.4.3 Developing New Criteria for Test Issues

Note: Skip to Step 3 of this section for "system supportability" issues.

"System operability" issues question the adequacy of system performance. There are two basic types of criteria that can apply to the question of performance: (1) the maximum time permitted for adequate performance, and (2) the minimum accuracy (i.e., the greatest amount of error) that is permitted for acceptable performance. Taken together, these two types constitute a unified criterion of acceptable overall performance of the system with regard to the test issue.

For the purpose of description, the process of identifying acceptable performance of a single performance "trial" (e.g., the time and accuracy of a single round of fire, etc.) is described separately from the process of identifying acceptable performance across several "trials" (e.g., time and percentage of hits in a burst of multiple rounds, etc.).

1. DETERMINE THE MINIMUM LEVEL OF ACCURACY THAT WILL BE ACCEPTABLE FOR ONE TRIAL OF SYSTEM PERFORMANCE.

This step is best done in several stages, namely:

A) Determine the dimensions that constitute performance accuracy. EXAMPLE: Target detection, target identification, distance from ammunition strikes to target center of mass, number of ammunition strikes per target, number of targets struck.

B) Identify the types of errors that can be made in each attribute (regardless of the magnitude of the error).
EXAMPLE: Presented target not detected, target misidentified, ammunition misses target center of mass, single target not hit on each of multiple attempts, each of multiple targets not hit.

C) Determine the greatest magnitude (size or maximum number) of error of each attribute that is still within the bounds of "acceptable" performance.
EXAMPLE: All targets presented must be detected. Target models may be misidentified (no other identification errors are permitted). Ammunition may miss center of mass of target by up to one meter. A target must be hit 3 times. Only one of four targets may be missed (not 3 times within one meter of center of mass).

2. DETERMINE THE MAXIMUM AMOUNT OF TIME PERMITTED FOR ADEQUATE PERFORMANCE OF ONE TRIAL.

The time of performance may be determined entirely from objective factors, such as the predicted "return fire" time of the threat system, or the time required to accomplish some machine-driven event. If such is the case, the maximum time for acceptable performance is determined analytically from these constraints. However, if the maximum performance time is subject to opinion, the following steps may be helpful:

A) Define the beginning and end points that delimit the event.

B) Consider various maximum permissible times for the event. For each maximum time considered, answer both of the following questions:

1) Will this performance time permit the successful completion of the overall mission(s) of the system?
2) Is this performance time possible within the defined accuracy and under this set of operational conditions?

The maximum time for which both of these questions can be answered "yes" is the time criterion for one performance trial.

EXAMPLE: The system will place at least 3 rounds of ammunition not more than 1 meter from the center of mass of at least 3 of 4 appropriately identified targets within 30 seconds.

3. DETERMINE THE CRITERION FOR MULTIPLE PERFORMANCE TRIALS.

Typically, the criterion of an issue is expressed in terms of performance across many trials. In this case, the criterion of "acceptable" performance may be stated as a percentage of successful trials, or as an average (median, or mode) performance level (with the associated variance). However the performance is to be calculated, the most useful criterion is a statement of combined time and accuracy of performance.

To determine the minimum acceptable percentage of successful trials (or minimum average performance), it is useful to consider the following questions:

A) Will the probability of success (or average level of performance) permit the successful completion of the overall mission of the system, and

B) Is this probability of success (or level of average performance) possible, given the definition of a single successful trial under particular operational conditions?

The minimum probability of success (or level of average performance) for which both of these questions can be answered "yes" is the criterion for this issue.

EXAMPLE: The system will place at least 3 rounds of ammunition not more than 1 meter from the center of mass of at least 3 of 4 appropriately identified targets within 30 seconds in 80 percent of the trials (or an 80 percent probability of occurrence).

4. DETERMINE THE CRITERIA FOR "SYSTEM SUPPORTABILITY" ISSUES.

For the purposes of testing the human performance aspects of an Army weapon system, the "system supportability" issues concern the adequacy of training, human factors engineering, personnel selection, and safety and health. In contrast to the "system operability" issues that directly question the adequacy of system performance, "system supportability" issues examine the contribution of the "supporting activities" to system performance. In effect, issues of supportability can be considered subsets of system operability issues. Thus, a supportability issue is deemed to be adequate to the extent that the system performance is acceptable. On the one hand, if the tested system performance meets the criterion of a system operability issue, it is reasonable to infer that the training, human factors engineering personnel selection, etc. were adequate in support of that performance. On the other hand, if the system performance does not pass the criterion, and if the system hardware did not malfunction, then it is reasonable to infer that one or more of these supporting activities was not adequate.

Theoretically, it may be possible to identify performance-based criteria for each of the primary supporting activities. However, for the purposes of test planning, it is generally sufficient to specify a priori criteria in a global and subjective manner. However, if the test subsequently identifies system performance that does not meet the criterion of a system supportability issue, it may be necessary to identify task-based criteria for some or all of the supporting activities that relate to that subcriterion performance.

2.5 Coordinating with Appropriate OT&E Documents

The issues, scope (conditions), and criteria developed in this chapter are inserted, initially, in the Independent Evaluation Plan (IEP) and subsequently carried over into the Test Design Plan (TDP). However, if issues, scope, and criteria have already been provided in the IEP, you may wish to review them in light of the HRTES procedures to assure that human resources will be tested adequately. These issues, etc., may then be inserted into the TDP.

3. PROCEDURE FOR PERFORMANCE TESTING

3.1 Overview

Typically, the next step in planning a test is to select the major tasks that are to be performed in the test. In the case of Operational Tests, these major tasks are sometimes defined in current system documents. The Human Resource Test and Evaluation System (HRTES) procedures are designed specifically to aid the test planner in identifying or developing the tasks that are related to the human resources of a weapon system, i.e., those tasks involved in operation of a weapon system that affect or are affected by human performance. Therefore, it is useful to review the system documents as a source for identifying those human performance related tasks that must be performed in the test of that weapon system.

This chapter provides guidelines for identifying tasks that will have human resources implications for test planning. Guidelines are also provided for developing tasks.

3.2 Identifying Tasks

The first step in planning performance testing is to identify the tasks that are required for operation and maintenance. In some cases, the documentation associated with the Army system will be sufficiently thorough to include all of the tasks required for each system operability issue. In these cases it simply is necessary to select the tasks from these documents, compare them with the operability issues, and verify that all tasks required by these issues are present. In other cases, there may be no tasks available, or only some tasks available, in the system documents. It then becomes necessary to develop additional tasks that will be required to test each operability issue. In the case of maintenance, scheduled maintenance tasks should be treated in the same manner as operational tasks. However, unscheduled maintenance tasks will have to be identified as they occur, or predicted in advance to the extent possible. Unscheduled maintenance tasks are to be measured only when a system breakdown occurs.

3.2.1 Selecting Tasks

1. IDENTIFY THE TASKS THAT ARE APPLICABLE TO THE ARMY SYSTEM BEING TESTED.

Depending on the stage of Army system development, some, most, or none of the tasks may already be given. At this point in test planning, it is necessary to determine which tasks are available and applicable to the system being tested. Relevant tasks may be contained in the documents associated with the development of the system, its previous testing, or the testing of similar systems.

2. COMPARE THE AVAILABLE TASKS WITH THE OPERABILITY ISSUES FOR THIS TEST TO DETERMINE THAT ALL REQUIRED TASKS ARE PRESENT.

There are two classes of test issues: (1) "system operability " issues that directly question the performance and maintenance of the system, and (2) "system supportability" issues that examine the various activities that contribute indirectly to system performance. The majority of system operability issues require some operator performance to be completed successfully. It is now necessary to compare the total list of tasks, for the system, with each operability issue to determine which tasks are required for the performance of each issue.

There are three classes of tasks: (1) Information input tasks (e.g., seeing and hearing), (2) Processing tasks (e.g., selecting, prioritizing, deciding, etc.), and (3) Output tasks (e.g., loading, aiming, firing, maneuvering, etc.). In most cases operability issues will require tasks from each of these classes.

While you are matching existing tasks to operability issues, keep these three categories of tasks in mind, and decide whether all required tasks are present for each issue. If you decide that some required tasks for an operability issue are absent, they will have to be developed later.

3.	DECIDE IF EACH AVAILABLE TASK IS EXPRESSED AT A MEASURABLE LEVEL. IF NOT, CONVERT EACH TASK TO A MEASURABLE LEVEL.

For a task to be measurable it should not be so small as to be impractical to measure (e.g., Turn on switch A, Read Speedometer, etc.), or so large as to be useless for analysis (e.g., Drive over sand, Engage target, etc.). If you have been given impractically small tasks, it is preferable to identify the function that these tasks are supposed to perform together, and to redefine the tasks into that larger function (e.g., Turn on switch A + Turn off switch B + Adjust knob C = measurable task--Zero the main gun). If you have been given excessively large tasks, it is preferable to divide them into their significant, measurable components (e.g., Drive over sand = measurable tasks--Start vehicle + accelerate + Maneuver using sharp turns + Back up vehicle + decelerate vehicle + Stop vehicle from high speed + Navigate vehicle in unknown terrain + Engage in high speed maneuver, etc.). You may find it useful to read through the tasks given in HRTES Supplement Section S3.1 as an aid to task conversion to a measurable level.

3.2.2 Developing Tasks

Note: Skip this section if all required tasks are available for all operability issues.

1.	DEVELOP ANY REMAINING TASKS THAT ARE REQUIRED FOR THE PERFORMANCE OF EACH OPERABILITY ISSUE.

In some cases, tasks required for the performance of each operability issue may not be available in the documents you have. To define the measures of performance and maintenance of the system, these additional tasks must be developed.

Procedures for developing tasks required for the performance of operability issues are given in HRTES - Supplement, Section S3.1.

3.2.3 Assigning Conditions to Tasks

1. IF THE SCOPE OF AN OPERATIONAL ISSUE CONTAINS MORE THAN ONE SET OF CONDITIONS, DETERMINE WHICH SET WILL APPLY TO THE TASKS REQUIRED BY THAT ISSUE.

In Chapter 2, guidelines were given for assigning scope (conditions) to each operational issue. It was stated that if there were mutually exclusive conditions for an issue (e.g., daylight and night) that it would be necessary to divide an issue's scope into sets of conditions that could be applied at the same time. If different sets of conditions are present in the scope of an operational issue, you will now have to decide whether all the required tasks (for that issue) will be measured under all or some of these sets. It is suggested that if tasks are to be measured under only one set of conditions, that it be the "worst case" set, since if the task is performed adequately under the worst case conditions, it is reasonable to expect the same (or better) performance under better conditions.

3.2.4 Developing Task Measures

1. DEFINE THE PERFORMANCE TIME MEASURE FOR EACH SELECTED TASK.

It is desirable to measure the performance time of each task selected. To do so, you will have to define three elements of each task:

(1) The unit of measurement (e.g., tenths of seconds, seconds, minutes, etc.).
(2) The beginning point of the task--That incident that causes performance timing to begin.
(3) The end point of the task--That incident that causes performance timing to cease.

In some cases the beginning point of a task will be the official start of a performance trial, whereas in some cases it will be the end point of the previous task. In come cases it will be necessary to give an artificial beginning point signal. The latter situation will exist when a number of tasks are being performed either simultaneously or sequentially without a break.

The end point of a task is often reasonably simple to define, but recognizing it in a field test may require that either a player or observer signal task completion.

Examples of tasks, beginning points, end points, and units of measurement follow:

TASK	BEGINNING POINT	END POINT	UNIT OF MEASUREMENT
Detect and Identify Target(s)	Introduction of First Target	Identification Signal	1/10 Sec.
Select and Order Targets	Identification of All Targets	Ordering Completion Signal	1/10 Sec.
Orient Weapon System	Start of Orienting	Completion of Orienting	Secs.
Determine Range of Target	Detection Signal	Range Call-Out	1/10
Shift to Second Target	Firing at First Target	Completion of Shift	Secs.

2. DEFINE THE PERFORMANCE ACCURACY MEASURE FOR EACH TASK SELECTED.

The simplest way of determining performance accuracy is to measure the types of errors that might reduce that accuracy. For example, if a task to be measured were "Identify enemy aircraft," the accuracy of its performance would be determined by (1) defining what constituted an identification error, (2) determining if it is necessary to record different types of identification errors or not, and (3) counting the errors in any given identification trial. The steps to define performance accuracy of a task follow:

(A) Determine the dimensions that constitute performance accuracy of the task.
EXAMPLE: Task--Target identification; Dimensions--Friendly vs. enemy status, Threat level, Class of target, Model of target, number of targets, etc.

(B) Determine the types of errors that can be made in each dimension.
EXAMPLE: Dimension--Friendly vs. enemy status; Errors--Friendly identified as enemy, Enemy identified as friendly, No friendly vs. enemy designation made, etc.

(C) Examine all possible error types for all dimensions of a task, and determine which types, if any, are acceptable or unacceptable.

(D) For those error types that are unacceptable, determine whether they can be collected under the general category of "Errors" (without differentiating them), or whether they must be differentiated by type.

(E) If types of errors have not been defined for unscheduled maintenance tasks by the time the field test occurs, determine accuracy by one or more of the following means:

(1) Expert opinion of observers.

(2) Ability of system to function following repair.

(3) Number of times maintenance task had to be performed before system could function.

In general, errors may be divided into the following categories: Errors of omission (something that should have been done was not done); Errors of commission (the wrong thing was done, or said, or identified); Sequence errors (the right things were done, but in the wrong order).

3. DETERMINE IF A SIGNIFICANT ACCIDENT MAY AFFECT PERFORMANCE MEASURE.

It is possible that a significant accident, or near-accident, may occur during a task trial. This information must be collected. However, it can be collected independently or as part of task performance measurement. It is suggested that any significant accident or near-accident that occurs should be associated with the task that was being performed and measured during its occurrence. In this way, it will be possible to use this accident data in evaluating task performance acceptability.

3.2.5 Logistics of Later Task Measurement

1. DETERMINE THE NUMBER OF PLAYERS AND TRIALS NECESSARY FOR RELIABLE MEASUREMENT.

This step is best accomplished by an analysis group. In the long run, assigning an appropriate number of players and trials per player is a necessity for producing meaningful test results. If you have insufficient numbers of players and trials, the OT evaluation will have little meaning for predicting real-world effectiveness of the system being tested. Analysis groups may be found at OTEA, TCATA, TRADOC, ARI, and other Army organizations.

HRTES Supplement, pages S3-41-48, includes a section on determining the number of players and trials, which you or the analysis group that you choose to utilize may find useful.

2. PLAN FOR THE LOGISTICAL CONSTRAINTS IMPOSED BY THE OT.

Measuring task-based performance requires considerable advanced logistical planning. It will be necessary to determine how the data is to be collected, and to plan for this collection. In general, task performance data will have to be collected by observers, instrumentation, or a combination of the two. Your decision as to the method of data collection will be based on various other logistical questions. If your function in OT planning includes making these logistical decisions about measurement, you may find the Worksheet on pages S3-49-56 of HRTES Supplement useful. Reading through it may trigger your thoughts and prove helpful at that level. Actually answering its questions may also prove to be a useful exercise leading to a better organized and planned OT.

3.2.6 Coordinating with Appropriate OT&E Document(s)

The output of this chapter consists of: tasks required for the performance of each operational issue, the conditions that will apply to each task, and the data requirements for the measurement of each task. This information should be included in the Independent Evaluation Plan and subsequent test planning documents. It may be useful to represent the tasks and their associated data requirements in the existing OTEA dendritic format.

4. COLLECTING ADDITIONAL DATA DURING OT

4.1 Overview

This chapter describes methods for planning the testing of training, human factors engineering (HFE), and safety during the field test. The major instruments for collecting these data are player and observer questionnaires. These questionnaires are based on identification of particularly difficult tasks, followed by ratings of the specific causes of difficulty. The questionnaires are augmented by critical incident report forms and safety checklists. Also provided (in HRTES-Supplement, Section S4.1) are player and observer checklists for the determination of specific causes of task difficulty.

1. DETERMINE THE TYPE OF PLAYER AND OBSERVER QUESTIONNAIRE TO USE IN THE FIELD TEST.

The alternatives are:

(a) Use the questionnaire contained at the end of this chapter (and checklists in S4.1 if desired);

(b) Have a questionnaire constructed that is specific to the system being tested.

Whichever alternative is selected, some form of questionnaire should be administered during the field test.

In addition to the player and observer questionnaire, this chapter includes:

(a) References to other methods of measuring training;
(b) Instruments for measuring system safety in detail;
(c) Description of HRTES' subjective measures of Human Factors Engineering;
(d) Description of method for flagging human factors inadequacies measured in the field test.

4.2 Training

1. DETERMINE WHICH STRATEGY YOU WISH TO EMPLOY FOR EVALUATING TRAINING.

Three strategies can be used to evaluate training during an OT:

(a) Evaluation of training activities independent of actual performance during the field test;
(b) Evaluation of training based on measurement of actual performance during the field test.
(c) Evaluation of training of those tasks that (1) were performed inadequately in the field test and (2) led to inadequate issue performance.

The first strategy assumes that one can determine the "goodness" of a training method by direct examination of that method (without field testing the performance being trained). It also assumes that it is desirable to understand the shortcomings of a training method even if the trainees are able to perform the trained tasks adequately.

The second strategy assumes that training is entirely responsible for the level of performance measured in the field test (assuming hardware does not malfunction), and that altering training can overcome all other system inadequacies.

The third strategy assumes that if trainees can perform measured tasks to criteria, the training of those tasks is acceptable and need not be examined further. It also assumes that if trainees could not perform one or more tasks to criteria, the training of those tasks might have been a cause and should be examined.

2. SELECT TRAINING MEASURES FOR THE TRAINING EVALUATION STRATEGY CHOSEN.

If you have chosen the first strategy (evaluation independent of performance), use of the following training measurement methodology should be considered: Guidelines for Conducting a Training Program Evaluation, ARI Working Paper FKFU 80-1, November 1979.

If you have chosen the second strategy (evaluation based entirely on performance measurement), you either may use HRTES without collecting any data on training method, or you should consider using the methods outlined in Training Effectiveness Analysis, Dept. of the Army, TRADOC Systems Analysis Activity.

If you have chosen the third strategy (training evaluation of inadequately performed tasks), you will have a choice of measuring training using the player and observer questionnaire, or using this questionnaire in conjunction with more specific training measures found in HRTES Supplement, Chapter S6.

These more detailed measures deal with:

(a) Training Time Allocation
(b) Adequacy of Practice Conditions
(c) Compatibility of Training Methods and Required Skills
(d) Adequacy of Personnel Who Trained the Task

These training measures in HRTES Supplement (Section S6-3 to S6-47) are directed to a specific task to be analyzed after the field test. However, you may decide that one or more of these measures requires data that must be scheduled for collection during, or immediately following, training. Therefore, it is suggested that you examine these training measures while planning the OT to determine if you may wish to utilize any of them for eventual analysis. If you decide to use any of these measures to analyze inadequately performed tasks (to determine if training is the cause) you may wish to schedule the collection of required data during the training segment of the OT.

4.3 Safety

1. DETERMINE IF YOU ARE SATISFIED WITH THE SAFETY QUESTION IN THE PLAYER AND OBSERVER QUESTIONNAIRE, OR IF YOU WISH TO COLLECT SAFETY DATA AT A GREATER LEVEL OF SPECIFICITY.

It is possible to measure system safety by keeping track of accidents (and near accidents) and by using the player and observer questionnaires. However, serious potential safety hazards may not lead to accidents in the field test or be noticed by players and observers. Therefore, use of the checklist (p. 4-15), or an equivalent checklist, is strongly urged and should be included in the Detailed Test Plan.

2. DETERMINE WHETHER YOU WISH TO COLLECT DATA ON ACCIDENTS AND NEAR ACCIDENTS DURING THE FIELD TEST

Significant accidents to personnel or equipment as a result of system operation or maintenance clearly affect system usability. Therefore, some version of a critical incident report is urged for inclusion in the Detailed Test Plan. An example of such a form may be found on page 4-11.

4.4 Human Factors Engineering

1. DETERMINE WHETHER THE QUESTIONNAIRE TO BE USED IN THE FIELD TEST IS ADEQUATE FOR COLLECTING HFE DATA FOR THE SPECIFIC SYSTEM BEING TESTED.

Examine the player and observer questionnaire and determine if it covers adequately the specific system being tested. If not, you may wish to add, modify, or eliminate questions. The center numbers and anchors of the questionnaire scales have been eliminated. This was done to avoid the usual respondant practice of automatically picking the center point of each scale. However, you may wish to introduce your own center points and anchors on these scales. It should be noted that objective HFE measures are listed in HRTES Supplement Chapter S6. These measures are used to analyze the causes of inadequate task performance.

4.5 Flagging Inadequacies

1. DETERMINE THE TYPE AND LEVEL OF INADEQUACY THAT WILL REQUIRE FLAGGING IN THE FIELD TEST.

There are four general sources of inadequacies that will be identified using HRTES:

(a) Questionnaires.
(b) Safety checklists.
(c) Critical incident reports.
(d) Causative analysis of inadequately performed tasks.

If the HRTES questionnaire is used, any system characteristic scale that receives a 50 or less should be flagged. The farther the scale is below 50, the stronger the reason for flagging it.

If HRTES safety checklist is used, any characteristic that receives a 50 or less should be flagged. The farther that characteristic is below 50, the stronger the reason for flagging it.

Any significant personnel or equipment accident should be flagged. The flagging of a near-accident depends upon the potential severity of that accident and the probability that it will occur in real system use. This probability may be estimated two ways: (1) by counting the number of near-accidents of a similar type, and (2) by questioning players and observers as to how close the near-accident came to actually taking place.

Inadequacies identified as a result of causative analysis should be flagged. This flagging procedure will be described in Chapter 6.

4.6 Coordinating with Appropriate OT&E Documents

The player and observer questionnaire, safety checklist, and critical incident form in this chapter are inserted in the Detailed Test Plan.

PLAYER AND OBSERVER QUESTIONNAIRE : 1

ISSUE: _____

STEP 1:

The tasks involved in performing this issue are listed below. Next to each task is a box that you should CHECK IF YOU HAD PROBLEMS PERFORMING THIS TASK. Leave the box blank if you did not experience any particular problems in performing the task.

TASK	CHECK IF PROBLEM
_____	☐
_____	☐
_____	☐
_____	☐
_____	☐
_____	☐
_____	☐
_____	☐
_____	☐
_____	☐
_____	☐
_____	☐
_____	☐
_____	☐

STEP 2:

Now that you have indicated which tasks you had problems performing, it is important to determine the source(s) of those problems.

NOTE: PLAYERS WILL FILL OUT THESE RATING SCALES ON THE BASIS OF THEIR OWN EXPERIENCES WITH TASK PERFORMANCE. OBSERVERS WILL FILL OUT THE RATING SCALES BASED ON THEIR IMPRESSION OF THE PLAYERS' PERFORMANCE.

PLAYER AND OBSERVER

QUESTIONNAIRE: 2

You checked tasks that caused problems. Fill in one of these Questionnaires for each task checked. First, write down the task in the space for it, below. Next, for this task rate each of the system characteristics given in the list below. If you are a player in this Test, rate all the system characteristics. If you are an observer, do not rate 19-22.

Your ratings are the percentage of your satisfaction with each system characteristic for performing this task. You can rate a system characteristic with any percentage from 0 to 100. The lower the percentage you give a characteristic, the more you are saying that it had a bad effect on the performance of this task. Write NA (Not Applicable) for system characteristics that do not apply.

TASK: _____

Not Satisfactory				Completely Satisfactory
0%	25%	50%	75%	100%

SYSTEM CHARACTERISTICS:	RATINGS:
1. Understandability of procedures required for task:	
2. Readability or hearability of displayed information used in task:	
3. Understandability of displayed information used in task:	
4. Usefulness of displayed information for this task:	
5. Ease of use of controls or other equipment used in task:	
6. Reachability of controls or other equipment used in task:	
7. Grouping (spatial configuration) of controls used in task:	
8. Ease of decision making required by task:	
9. Visibility of the target or other parts of the environment:	
10. Ability to see parts of work station (controls, displays, etc.):	
11. Noise level where you were located:	
12. Effects of motion on performance of this task:	
13. Effects of ventilation on performance of the task:	
14. Effects of workspace temperature on performance of the task:	

TASK: _____

Not Satisfactory				Completely Satisfactory
0%	25%	50%	75%	100%

SYSTEM CHARACTERISTICS: RATINGS:

15. Effects of the dimensions (size and layout) of the workspace: [ ]
16. Effects of seating design on performance of the task: [ ]
17. Workload of task plus any other tasks performed at same time: [ ]
18. Effects of design for safety on performance of the task: [ ]

THE FOLLOWING SCALES ARE TO BE RATED ONLY BY PLAYERS:

SYSTEM CHARACTERISTICS RATINGS:

19. Training time for this task: [ ]
20. Methods used to train this task: [ ]
21. Adequacy of the length/type of practice of this task: [ ]
22. Adequacy of the trainer in training this task: [ ]

OTHER - Please describe other system characteristics or other factors that you think produced problems in task performance.

CRITICAL INCIDENT REPORT

TASK

WHAT HAPPENED?

WHAT PRODUCED THIS PROBLEM?

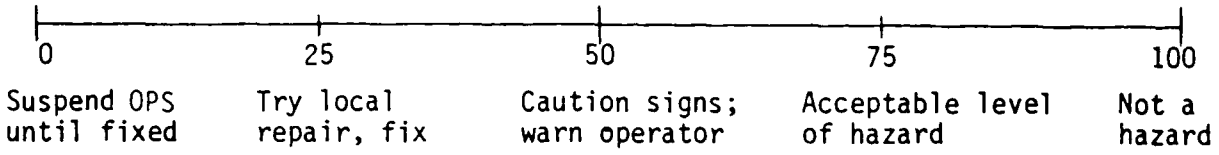
HOW DID YOU DISCOVER THIS PROBLEM?

HOW DID YOU, OR WOULD YOU, SOLVE THIS PROBLEM?

WHAT DID IT, OR COULD IT, HAVE CAUSED?

- ☐ PLAYER
- ☐ OBSERVER

SAFETY CHECKLIST



CHARACTERISTICS	SAFETY RATING
SLIPPERY WALKING SURFACE	
SLIPPERY CLIMBING SURFACE	
CLIMBING SURFACE WITHOUT ADEQUATE FOOTHOLDS	
LIFTING/CLIMBING SURFACE WITHOUT ADEQUATE HANDHOLDS	
INADEQUATE GUARDRAILS/SHIELDING	
DANGEROUS HARDWARE/SOFTWARE CONDITION NOT ADEQUATELY SIGNALED	
DANGEROUS ENVIRONMENTAL CONDITION NOT ADEQUATELY SIGNALED	
DANGEROUS TACTICAL CONDITION NOT ADEQUATELY SIGNALED	
DANGEROUSLY HIGH AIR TEMPERATURE	
DANGEROUSLY LOW AIR FLOW/TIME	
EXCESSIVE INTERNAL COMBUSTION OR GUNFIRE PRODUCT LEVEL	
CONTROL DANGEROUSLY HARD TO MANIPULATE/REACH	
DISPLAY DANGEROUSLY HARD TO READ/UNDERSTAND	
ELEMENTS IN ENVIRONMENT DANGEROUSLY HARD TO SEE	
OTHER:	

CHARACTERISTICS	SAFETY RATING
SHARP EDGED OBJECT	
POINTED OBJECT	
SNAGGING OBJECT	
SMALL DIAMETER PROJECTION	
DANGEROUSLY INADEQUATE HEAD CLEARANCE	
EXPOSED EXCESSIVELY HOT MATERIAL	
EXPOSED EXCESSIVELY COLD MATERIAL	
EXPOSED SOURCE OF ELECTRIC SHOCK	
EXPOSED MACHINERY IN MOTION NOT ADEQUATELY HIGHLIGHTED	
TOXIC MATERIAL/RADIATION CONTACTABLE	
NOXIOUS MATERIAL CONTACTABLE	
SOUND PRESSURE AT DANGEROUS LEVEL	
VIBRATION AT DANGEROUS AMPLITUDE/FREQUENCY	
DANGEROUSLY INADEQUATE ILLUMINATION OF POTENTIAL ACCIDENT SITE	
DANGEROUSLY EXCESSIVE ILLUMINATION	
INADEQUATE EQUIPMENT ANCHORING	
INADEQUATE PERSONNEL RESTRAINT	
INADEQUATE EQUIPMENT PADDING	

5. EVALUATION

5.1 Overview

This chapter discusses methods for identifying the task(s) that led to unsuccessful performance of an operational issue. It is based on the following steps:

(a) Determining which operational issues did not meet their previously defined criteria;
(b) Determining which tasks made up the issues that did not meet their criteria;
(c) Identifying the conditions that were present when the tasks were measured;
(d) Determining the types of data that were taken for each task;
(e) Developing single trial criteria for each of these tasks;
(f) Developing multiple trial criteria for those tasks that require them;
(g) Identifying the task(s) that led to inadequate issue performance by comparing task data with criteria.

5.2 Identifying Tasks that Led to Issue Failure

1. DETERMINE WHICH OPERATIONAL ISSUES DID NOT MEET THEIR MULTIPLE TRIAL CRITERIA.

This is done by comparing the field test data for each operational issue with the multiple trial criterion for each issue. If the measure of an operational issue falls below its multiple-trial criterion, that issue has problems and is a candidate for further analysis.

2. FOR THOSE ISSUES THAT DID NOT MEET THEIR MULTIPLE-TRIAL CRITERIA, DETERMINE ON WHICH TRIALS THE ISSUE PERFORMANCE DID NOT MEET ITS SINGLE-TRIAL CRITERION.

This is done by comparing the field test data for each trial of each operational issue that failed with the single-trial criterion for that issue. If the trial measures for a particular issue did not meet the single-trial criterion for that issue, this indicates problems in issue performance during that trial, and should be investigated further.

3. FOR OPERATIONAL ISSUE TRIALS THAT DID NOT MEET THEIR SINGLE TRIAL CRITERION, DETERMINE WHICH TASKS WERE MEASURED.

In the preparation of the Outline Test Plan and Detailed Test Plan, the tasks required for the performance of each operational issue were defined through the HRTES procedure. Further, if HRTES had been used, some or all these tasks were measured in the field test. Now, you have to associate the issue trials that fell below their single-trial criteria with their associated tasks that were performed and measured.

4. DETERMINE THE TYPES OF MEASURES USED FOR EACH OF THE TASKS IN QUESTION.

Each task was measured in terms of its performance time and/or performance accuracy. Performance time was defined in terms of that task's beginning and end points and the units of time measurement (tenths of seconds, seconds, minutes, etc.). Accuracy was defined in terms of the types of performance errors to be measured.

5. IDENTIFY THE CONDITIONS UNDER WHICH THE ISSUE TRIALS FELL BELOW THEIR SINGLE-TRIAL CRITERION.

6. DETERMINE THE MINIMUM ACCEPTABLE LEVEL OF ACCURACY FOR ONE TRIAL OF TASK PERFORMANCE.

This step is best done in several stages:

(a) Identify the types of errors that can be made in task performance. The types of errors will have been defined prior to the field test.

(b) Determine the greatest magnitude (size or maximum number) of error that is still within the bounds of "acceptable" task performance under the actual conditions during measurement.

NOTE: The criterion for all tasks (of a particular issue) taken together cannot exceed the criterion of the issue itself.

7. DETERMINE THE MAXIMUM AMOUNT OF TIME PERMITTED FOR ADEQUATE PERFORMANCE OF ONE TRIAL OF THE TASK.

Performance time for a task may be determined entirely from objective factors, such as the predicted "return fire" time of the threat system, or the time required to accomplish some machine-driven event. If such is the case, the maximum time for acceptable performance is determined analytically from these constraints. However, if the maximum performance time for a task is subject to opinion, the following steps may be helpful:

(a) Define the beginning and end points that delimit the task. These should be available in the Detailed Test Plan.

(b) Consider various maximum permissible times for the task. For each maximum time considered, answer both of the following questions:

(1) Will this performance time permit the successful completion of the parent issue, and
(2) Is this performance time possible within the defined accuracy requirements and under this task's set of operational conditions?

The maximum time for which both of these questions can be answered "yes" should be the time criterion for one performance trial.

NOTE: As was the case for accuracy, time criteria for all tasks (of a particular issue) taken together cannot exceed the time criterion for the issue itself.

8. FOR EACH ISSUE TRIAL THAT FAILED TO MEET ITS SINGLE-TRIAL CRITERION, DETERMINE WHICH TASKS HAD TO BE PERFORMED MORE THAN ONCE.

If you determine that any tasks had to be performed more than once in the trial of their parent issue, you must develop a multiple performance criterion for each of those tasks.

9. DEVELOP MULTIPLE PERFORMANCE CRITERIA FOR TASKS THAT HAD TO BE PERFORMED MORE THAN ONCE DURING AN ISSUE TRIAL.

The multiple performance task criterion may be in the form of a percentage of successful performances, or an average (median or mode) performance level, with associated variance. The most useful criterion is usually a statement of combined time and accuracy of performance.

To determine the minimum acceptable percentage of successful task performances (or minimum average performance), it is useful to ask:

(a) Will the probability of task success (or average level of performance) permit the successful completion of the parent issue, and

(b) Is this probability of success (or level of average performance) possible, given the definition of a single successful task trial under particular operational conditions?

The minimum probability of success (or level of average performance) for which both of these questions can be answered "yes" is the multiple performance criterion for this task.

10. IDENTIFY THE TASK(S) THAT LED TO ISSUE FAILURE.

For each issue trial that fell below its criterion, compare task data with task criteria. Those tasks that fell below their criteria (single trial criterion, of--if repeated tasks--multiple performance criterion) are the tasks that led to issue trial failure. All such tasks are significant and should be considered for further analysis (see Chapter 6). The larger the percentage of issue trials during which a given task failed, the more significant that task. The more significant a task, the greater the urgency that it be further analyzed.

5.3 Coordinating with Appropriate OT&E Documents

The following information, developed in this chapter, should be inserted in the Independent Evaluation Report:

(a) Issues that passed their criterion and those that fell below criterion;

(b) The conditions that applied to both classes of operational issues.

(c) Tasks that led to operational issue failure, the criteria of those tasks, and their performance data. Particular emphasis should be given to those tasks that are associated most often with issue trials that failed to meet their criteria.

6. ANALYSIS

6.1 Overview

This chapter describes strategies for determining the reasons for inadequate performance of those tasks that were identified in Chapter 5. There are three classes of human factors measures that should aid in making this determination. These are measures of training, human factors engineering (HFE), and personnel characteristics of the field test players. Two strategies are described for taking such human factors measures of tasks: using data collected in the player and observer questionnaires that were completed during the field test, and having human factors area specialists take detailed measures. HRTES Supplement (Chapter S6) contains a comprehensive list of such measures, a method for selecting such detailed measures for each task, and a method for reducing their data into an easily understood format for analysis.

6.2 Determining Human Factors Causes for Task Failure

As a result of developing task criteria for the tasks that are components of inadequately performed operational issues (in Chapter 5), you know which tasks were associated with inadequate issue performance. You now must determine how best to analyze each of these tasks.

1. DETERMINE THE STRATEGY TO BE USED FOR ANALYZING TASKS.

There are two strategies for analyzing tasks: (1) Using data from player and observer questionnaires (collected during the field test); (2) Using detailed measures of human factors areas (taken by specialists).

The first strategy will provide a speedy analysis and will not require the intervention of training, HFE, or personnel specialists. However, the data it produces will be based entirely on the opinions of players and observers in the field test. The second strategy will be based on detailed human factors measures (both objective and subjective) taken by area specialists. When specific measures are not available in the time permitted, they will be replaced by data from the player and observer questionnaires. This strategy may produce conclusions of higher validity than the first strategy. However, it requires specialist intervention and will take longer to implement than the first strategy.

Consider the following elements when deciding which strategy to use:

(a) The criticality of the system that was tested in the OT;
(b) The eventual expense of retrofitting the system after production;
(c) The probability of producing change in system design at this stage of the OT&E cycle;
(d) The criticality of the operational issue that was affected by inadequate task performance;
(e) The percentage of inadequate issue trials that were affected by a given, inadequately performed task;
(f) The amount of real-time available for analysis;
(g) The amount of professional man-hours available for analysis.

NOTE: If a comprehensive test and evaluation of training methodology (in the OT) has already been conducted for this system, a combination of the first and second strategies should be considered for training measures.

2. DETERMINE WHICH CLASSES OF HUMAN FACTORS MEASURES TO APPLY TO THE TASK BEING ANALYZED.

There are three classes of human factors measures that can be used to analyze a task: (1) training measures, (2) HFE measures, and (3) measures of personnel characteristics. The determination of the classes of human factors measures to take applies to either strategy that you selected in the preceding step.

(a) Training--Analysis should include possible training causes as measured by strategies 1 or 2 or a comprehensive training methodology evaluation,

(b) HFE--Analysis should include possible HFE causes as measured by strategies 1 or 2 or some other HFE system.

(c) Personnel Characteristics--If you are sure that the players in the OT were a representative sample of the actual users of the system, personnel characteristics need not be measured. However, if you have a good reason for suspecting that the players were not a representative sample, you should consider measuring personnel characteristics. If you have selected strategy 1, personnel characteristic measures will not be available. If you require personnel characteristic measures, you will have to select strategy 2 to obtain them.

HRTES Chapter S6 contains sections for the use of specialists in taking measures of training, HFE, and personnel selection. If you have decided on the second strategy for one or more classes of human factors measures,

it may prove useful to present the appropriate section(s) of Chapter S6 to the specialist(s) who will be taking these measures.

3. DETERMINE THE HUMAN FACTORS CAUSE(S) FOR THE FAILURE OF TASK PERFORMANCE.

Strategy 1 Data--Strategy one data consist of rating scale values ranging from 0 to 100. These values apply to system characteristics that apply to each task that a player or observer considered difficult. The following steps apply to each task being analyzed:

(a) Retrieve the player and observer questionnaires for the task.

(b) If possible, retrieve those questionnaires completed by the player(s) whose performance fell below criterion and the observers that observed that performance.

(c) Examine the values assigned to the system characteristic scales in those questionnaires. Any value significantly less than 100 indicates the presence of a potential problem in the system design for that task. The lower the value, the greater the problem.

Strategy 2 Data--This strategy is based on measures taken by specialists in the various human factors areas. The nature of the data you receive will vary according to the measurement system used. If the measurement system described in HRTES Chapter S6 was used, you will obtain the following types of data:

(a) Detailed measures for each task.

(b) Information as to whether measures met their criteria.

(c) Condensed, hierarchically arranged summations of each class (training, HFE, personnel characteristics) of measures. In these summations (called "Summary Worksheets") individual problems causing task difficulty will be listed. In addition, figures of merit at various levels, up to the level of training, HFE, and personnel characteristics, will be listed for each analyzed task. These figures of merit are composed of specific measures that were taken either by specialists or from questionnaires. The lower the value of a particular figure of merit, the greater the problem associated with that figure.

6.3 Coordinating With Appropriate OT&E Documents

The summation of questionnaire values, human factors measures and their figures of merit, and the resulting analyses should be inserted in the Independent Evaluation Report.

7. SUMMARY

You have now reached the end of HRTES Test Procedures and its accompanying Supplement. They were designed to aid you to integrate the significant aspects of human performance into the test and evaluation cycle. In summary, HRTES included methods for aiding in the selection or development of the following elements that are used in test and evaluation:

(1) Test Issues
(2) Scope
(3) Issue Criteria
(4) Tasks
(5) Task Measures
(6) Logistics of Task Measures
(7) Number of Players and Trials
(8) Attitudinal Measures (Performance)
(9) Attitudinal Measures (Safety)
(10) Issue Evaluation
(11) Task Criteria
(12) Task Evaluation
(13) Human Factors Measures
(14) Causal Analysis of Inadequate Task Performance

The integration of human performance into overall system test and evaluation is a process that leads to the goal of enhanced systems acquisition.